View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Carolina Digital Repository

Scaling Down Transaction Log Analysis: A Study of OPAC Usage at a Small Academic Library

by

Mitake Holloman

A Master's paper submitted to the faculty of the School of Information and Library Science of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in Library Science.

Chapel Hill, North Carolina

July, 1999

Approved by:

_____

Advisor

Mitake Holloman. Scaling Down Transaction Log Analysis: A Study of OPAC Usage at a Small Academic Library. A Master's paper for the M.S. in L.S. degree. July, 1999. 54 pages. Advisor: Barbara M. Wildemuth.

This study looks at a month's worth of On-line Public Access Catalog (OPAC) transaction logs from the Broyhill Learning Resource Center at Caldwell Community College & Technical Institute in western North Carolina. The study was conducted to look at usage of an OPAC in a smaller library setting; to understand any problems occurring in the OPAC that suggest needed changes in interface design and/or bibliographic instruction; and to identify sets of data that may be useful for the library managers to review on a regular basis.

Over 83% of the 5,479 searches performed were executed as subject searches. The heaviest usage of the catalog occurred when classes came in to use the library. About 35% of the searches resulted in zero records being retrieved. Over 60% of the zero-hit searches were caused by problems with the controlled vocabulary. Information from transaction logs should be useful for the library to keep track of system usage, search failure rates, advanced feature usage, and common search strings.

Headings:

Junior and community college libraries -- Online catalogs

Junior and community college libraries -- North Carolina

End user searching -- Evaluation

Use studies -- Online Catalogs

Table of Contents

List of Tables

List of Figures

Introduction

On-line Public Access Catalog (OPAC) transaction logs unobtrusively and anonymously record the text of searches performed during a specified period. They can also record information such as the number of database hits retrieved by a search, and if title level displays were viewed or not. The first phase of transaction log analysis (mid 1960's to late 1970's) focused mostly on system performance instead of user interactions. Since the early 1980s, when on-line catalogs started to become available, a large number of OPAC transaction log studies have focused on many different aspects of user behavior and interactions with OPAC systems.

The focus of these studies is highly varied. One review of the field (Peters, 1993) breaks the types of analysis done on public use of deployed systems into 19 categories. These categories include: commands, response time, session lengths; chains of commands and state to state transitions; inter-mode analyses -- menu and command; intersite and interinstitutional analyses -- use of union scoping; intersystem analyses; errors, zero-hits, missed opportunities, failures, and their causes; on-line help and other instructional opportunities; analyses of specific search states; extent of match studies; other access points and advanced features; printing and downloading behavior; user persistence; quitting behavior and situations; analyses of special public user groups; longitudinal studies and variations over time; replications of previous transaction log analysis (TLA) research; TLA combined with other methods; analyses of other databases, and analyses of environmental factors. Stud-

ies can also be found on staff use of deployed systems and general use of experimental systems.

Most studies have focused on searching behaviors of users in large academic libraries. This study seeks to apply transaction log analysis to a smaller setting and determine the usefulness of this method as a long term management tool.

Literature Review

Analyses of OPAC Use

In trying to understand the effectiveness, ease of use, and user satisfaction with an OPAC it is generally useful to look at the searching behaviors of the users and analyze the search failures that are generated. Although their specific goals have differed, a number of studies have reported on general searching behaviors and reasons for search failures in an effort to provide a baseline of information. These studies include studies such as Peters' (1989) and Zink's (1991) failure analyses, Hunter's (1991) and Wallace's (1993) look at the implications for bibliographic instruction (BI), Wyly's (1996) identification of the access points of records that the users found interesting, and Atlas' (1997) examination of changes in searching after the placement of documentation by the terminals.

The search behaviors described by these studies suggest a wide variety in choice of access points across user populations and systems. In these six studies the usage of author as an access point varied from 13% in the Zink study to 22% in the Peters study. Usage of titles varied from 17% in the pre-documentation study by Atlas to 34% in the Peters study, and subjects were used in only 26% of the searches in the Wyly study but in 52% of the searches in the Hunter study. The zero-hit rates in the studies varied from 10.4% in the Wallace study to 54.2% in the Hunter study (Peters, 1989; Zink, 1991; Hunter, 1991; Wallace, 1993; Wyly, 1996; and Atlas, 1997). In his review of transaction log studies, Peters (1993) notes that there is an amazing

diversity among these studies, even those purporting to study the same questions. The variances are likely due to the variety of features available across systems and the variety of user expertise and understanding that is available in the studied populations.

It has also been noted that the use of subject searching with the traditional Library of Congress Subject Headings (LCSH) has declined over the years as other options have proven to be more successful. Larson (1991) documented a sharp decline in the use of subject searching with the MELVYL system used by the University of California over a period of about six years. In 1982 subject searching accounted for 70-80% of the user searches, while known item searches (those on the title and author indexes) were about 20-30% of the searches. By 1988 known item and subject searching were about equal in popularity. Larson attributes this trend to two factors, the relatively high failure rate for subject searches due to the need for knowledge of how LCSHs work, and the increasing likelihood of information overload as the database increases and returns datasets that are unmanageably large.

Many of the reasons for search failures appear consistently throughout the literature. In fact, they are consistent enough that these errors are well described in the practical literature of library and information science as well as the research literature. The most common reasons for user errors in OPAC searching include input error such as improper truncation or spelling errors, inability to understand Boolean logic, poor choice of search terms, navigation problems through multi-database systems, unwillingness to read the help files or documentation, and conceptual

errors such as not understanding the scope of the catalog or the strengths and limitations of the system (Tenopir, 1997).

Although typographical and spelling errors are very common (they account for about 25% of the errors in the Zink (1991) study and about 15% in the Peters (1989) and Hunter (1991) studies) it is not obvious that just fixing these errors will solve the problems of the user. Drabenstott (1996) found that only 6% of user queries had a spelling error that, if corrected, could be found in the database. More often the user made other errors as well that may have been more detrimental to the search. This suggests, that while spelling is a problem for many users, user who misspell their search terms will often have other hurdles to clear before getting results from an OPAC.

The nature of the hurdles that a user may have to clear is generally dependent on the knowledge that they bring to the searching situation. Borgman (1989) found that technical aptitude, comfort with math and computer science, and logic skills were strongly correlated with success in information retrieval tasks. She also breaks down the knowledge that a user must have to be effective and efficient at searching into technical, semantic, and conceptual areas (Borgman, 1996). Technical knowledge often falls into the areas of computer literacy, including such things as layout of the keyboard and how to operate the input devices, and an understanding of the system syntax and how to manipulate the commands. Semantic knowledge involves an understanding of the actions to take to complete a search, the access points to use, the appropriate search terms, boolean logic, and the organization of files within the system. Conceptual knowledge demands an understanding of the

search goal and the ability to identify when it has been reached. In another study looking into what kinds of knowledge are needed for success in information retrieval tasks, Allen (1991) suggests that, while catalog knowledge, knowing how to manipulate the search system, and cognitive ability are important for searching, topic knowledge may not be as important for databases such as on-line catalogs because the depth of information is not available. The findings of his study showed that, when working with information surrogates such as catalog records, there are often not enough records or depth to the records to help someone who is simply more topic knowledgeable to distinguish the most important item.

That said, there is a fair amount of difference in the ways different user populations approach the OPAC. In studying users of the MELVYL system at the University of California, Ferl (1996) noted that undergraduate users performed subject searches about 51% of the time whereas faculty, graduate students, and library staff only used subject searching 15-20% of the time. This suggests that as user knowledge in topic areas and/or searching skill increases users change their search strategies.

Since users of smaller library systems such as community colleges and rural libraries can be expected to have knowledge and skills that are lower than university undergraduates suggests that the search patterns for these institutions may be different than for the large four year universities with graduate programs which are most commonly studied. Very few studies have been done on smaller libraries. The Council on Library Resources (CLR) study undertaken in the early 1980s looked at a variety of libraries including a community college system in Santa Clara, California.

The Mission and West Valley Community Colleges served about 19,000 students and had about 79,000 records in their database. This is a sharp contrast from Mankato State University (Minnesota) which was also in the study and served about 15,000 patrons with 311,000 items (almost an order of magnitude difference in collection size). Both the community colleges and Mankato State were observed for a month in the study. The community colleges accumulated about 26,000 OPAC transactions; Mankato State accumulated about 240,000 OPAC transactions. Due to the limitations of computing power at that time, neither of these organizations were included in the transaction log analysis portion of the study; they were only involved in the questionnaires. In the study the community colleges each had about 64% of surveyed searchers report that they were looking for a topic, 27% looking for a specific book, and 9% for a specific author. Mankato State had 76.5% report that they were looking for a topic, 21.5% for a specific book, and about 2% looking for a specific author. Among the 14 libraries in the study 9 reported a rate of topic searching that was lower than the two community colleges (minimum 37.1% at Northwestern), and 3 reported a rate for subject searching that was higher (maximum 76.5% at Mankato State) (Matthews, 1982; and Matthews, 1984).

Unfortunately it is often difficult to compare the results of the studies and apply them to systems other than the one studied, especially when trying to identify problems that are occurring for patrons that may have a different level of experience and expertise. The information available from an OPAC's transaction logs will vary from one system to another. Also, an OPAC will typically have its own specific screen layouts, special features, searching aids, and index definitions. Attempts to

compare transaction log analyses have generally been limited to comparisons of databases on the same platform at different locations and general statements about what indexes patrons are using to search, the fact that they make lots of typographical mistakes, and that they don't understand controlled vocabulary. Even these statements can be hard to compare because of variations in how different OPACs have developed their interfaces and features. Furthermore, while the collection of transaction log data is fairly easy, the analysis of that data is not. The studies listed above and others like them represent many hours worth of work and require a large commitment of resources to undertake. This is generally not feasible for smaller community college and public libraries.

Applying TLA to Management Decision Making

In terms of gathering management information, transaction logs can be used to determine when high demand periods occur, the locations from which the users are accessing the system, the impact of self-serve services, trends in collection usage, and the efficacy of an attempted system enhancement or services tangential to the system itself such as bibliographic instruction (BI). Information about peak demand periods can be useful in scheduling of system maintenance and service hours. Location information can reveal areas that might have queuing problems or need special targeting, such as more training at a branch library. The collection usage trends can be used in determining areas of needed focus for collection development. The impact and efficacy information can, in general, be used to justify the allocation of resources for any given project or highlight areas that could be more effective. "Transaction log analysis serves as a reality check, rude awakening, or

cold slap in the face of all types and levels of management related in some way to the use of an automated library system" (Peters, 1996, p. 24). The trick to making effective use of the information available for management purposes is an effective analysis of the library's needs and the development of reports that categorize the information into digestible chunks (Sandore, 1993; Peters, 1996).

Transaction log analysis provides an unobtrusive observation method with a focus on user behavior rather on user attitudes. It can provide a valuable set of information about how the library system is being used absent of user bias and, when combined with other research methods, can paint a fuller picture of many aspects of library usage. This information is no less useful for smaller libraries that large academic institutions.

Research Questions

The goals of this study are to look at usage of an OPAC in a smaller library setting; to understand any problems occurring in the OPAC that suggest needed changes in interface design and/or bibliographic instruction; and to identify sets of data that may be useful for the library managers to review on a regular basis.

As a result the specific questions that will be explored include:

1. How many searches are being performed?
2. What kinds of searches are being performed?
3. Is there a pattern of times, days, weeks that are more popular than others?
4. How many of the searches are failing (zero-hit and too many hits)?
5. What are the reasons searches are failing?
6. How would the searches be affected by changing the index, to search all the indexes?
7. How frequently are users using advanced search features of the system?
8. What improvements to the system are suggested by the current usage patterns?
9. What tools are needed to monitor the system over time?

Environment

The library

Caldwell Community College and Technical Institute (CCC&TI) serves about 2800 students per semester in its mission to "provide accessible, quality, educational instruction in college transfer, literacy, and occupational programs to individuals who seek to enhance their employment opportunities" (Caldwell..., n.d.). The college also runs a variety of economic development programs for local businesses and agencies. There are approximately 30 major programs offered at the college. They include accounting, HVAC technology, auto body and auto repair technologies, business administration, cosmetology, early childhood associate, law enforcement, nursing, office systems technology, pharmacy technology, and various medical technologies. Other programs offered by CCC&TI include an adult high school program and a college parallel program that allows students to take the equivalent of the first two years of undergraduate studies. The college is located in western North Carolina on two campuses. The main campus is in Hudson (located between Lenoir and Hickory) and the other is in Boone. Students are expected to commute to the school and many classes are offered at night and some on the weekends.

The Learning Resource Center (LRC) provides a support collection for degree programs offered by the college. The LRC has a full time staff of four, about 48,500 monograph titles, about 250 serial titles, and an annual circulation of 18,500.

The OPAC

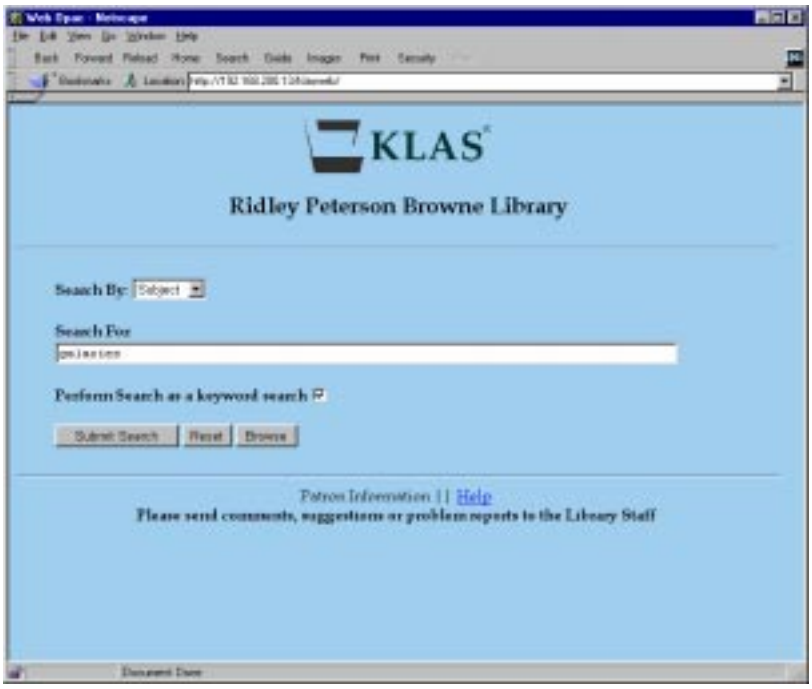
The Broyhill Learning Resource Center (LRC) at CCC&TI uses the Keystone Library Automation System (KLAS) to manage their library operations. Their OPAC is fully integrated into the rest of the management system and displays the current status and availability of all of the Center's catalogued materials. The OPAC is a Web based OPAC, meaning that it runs in a standard web browser such as Internet Explorer or Netscape instead of on character terminals. Currently the OPAC is only available on the college's local area network and is mostly accessed through one of the 22 workstations in the LRC. Also about 25-30 instructors have requested a quick access icon from their campus workstations.

Upon beginning the search process users are greeted with the search screen shown in Figure 1. Users are able to perform keyword and "begins" searches on three indexes: author, title, and subject. "Begins" searches match the string entered to the beginning of a record. For example, a "begins" search on the author index for "Adam" would retrieve all of the authors whose last name begins with the letters A-D-A-M, including Adams, Adamson, Adamy, etc. The same search done as a key-word search would return all authors who had "Adam" as a name but would not retrieve the last names such as Adams. A keyword search with multiple terms will return items containing all of the entered words any where in the fields specified by the index. Searches are case insensitive, strip out punctuation, and ignore initial articles.

The indexes currently available at CCC&TI are Author, Title, and Subject. The library has choosen for the author index to search against the 100, 110, 111, 700,

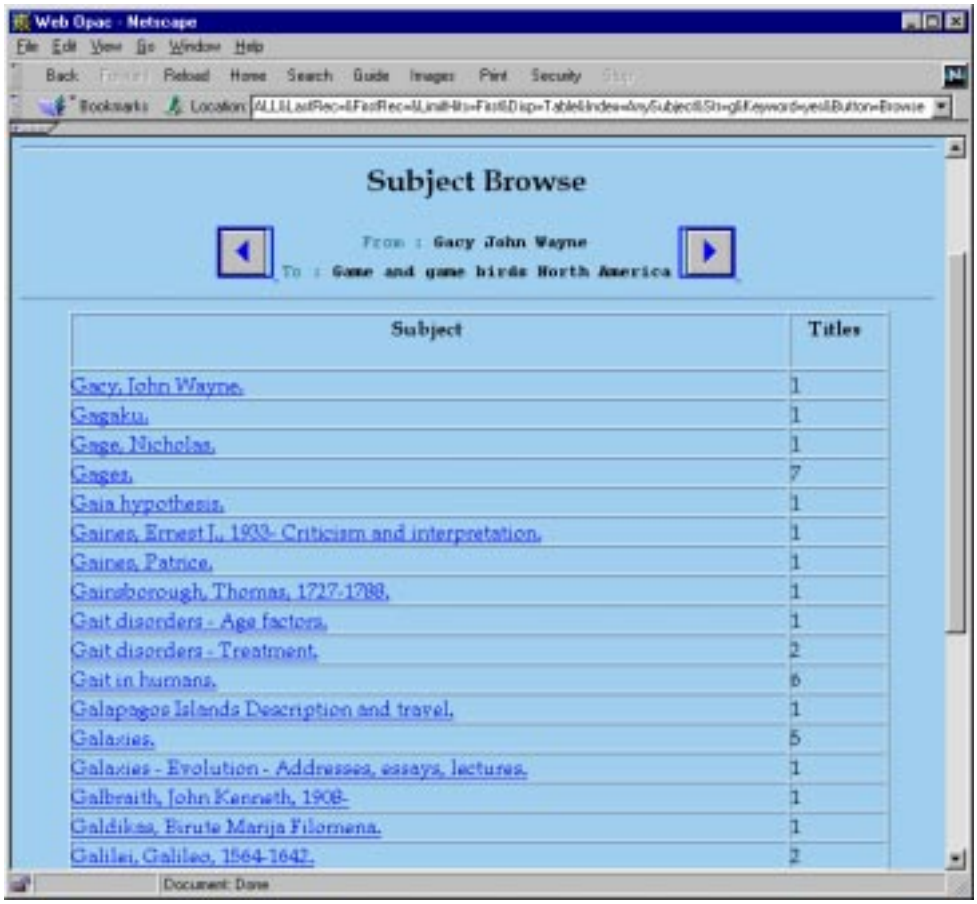
710, 711, and 505 MARC tags. The title index searches on the 245ab, 440ab, 740ab, 242ab, 246ab, 700t, and 505 tags. The subject index searches on the 600, 610, 611, 630, 650, 651 tags.

Figure 1: Search Screen



Users also have the ability to browse the indexes; the browse list indicates the number of records matching a particular index entry. Figure 2 shows a browse screen that would result from the user selecting the browse button in Figure 1. Since the user had entered “galaxies” as a subject search, trems from the subject index in the alphabetical vicintity of “galaxies” are displayed. The number of titles in the collection using a particular subject are shown in the right column. From this screen the user can perform a search on a given subject by clicking on the desired subject term, page through the lists of subject terms, or go back to the search screen using the information they have gained about the library’s holdings to formulate their search.

Figure 2: Browse Screen

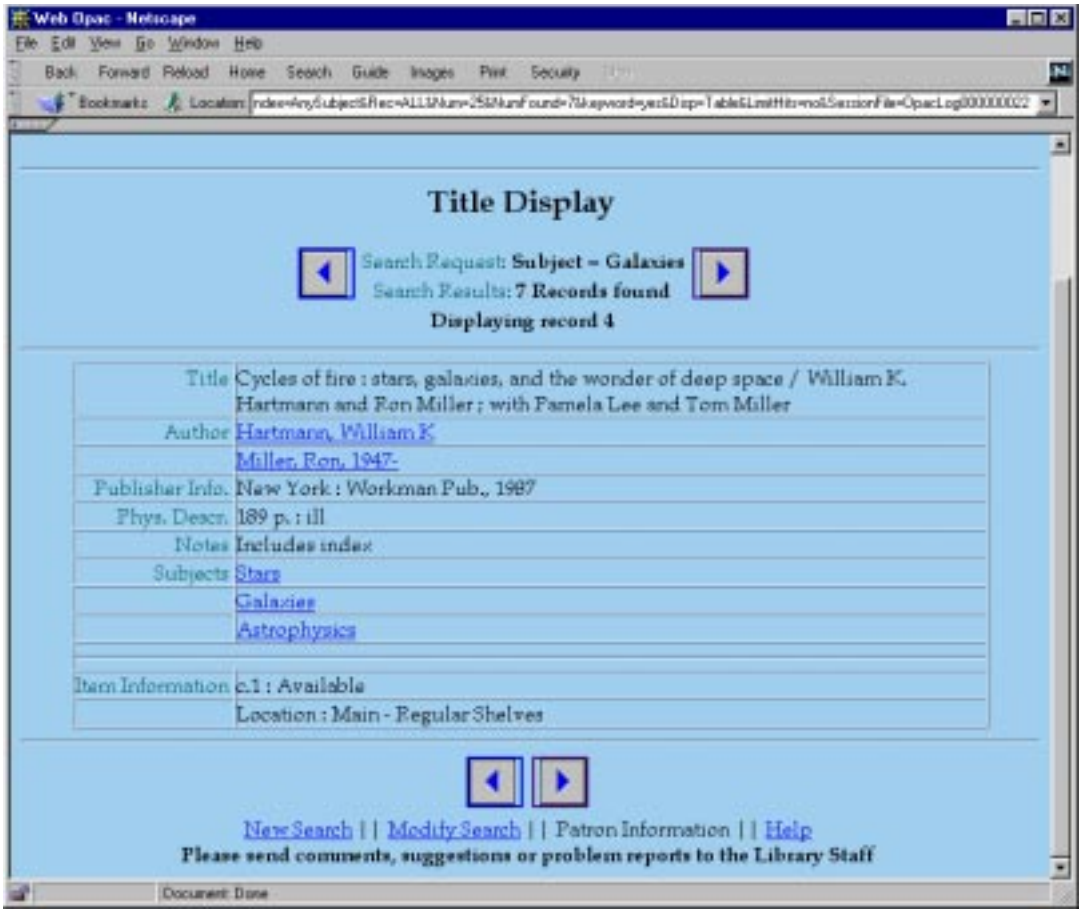


Searches that retrieve hits return the Search Results display (Figure 3) which displays up to 25 hits per page. The user can use the next and previous buttons to navigate among pages of search results, click on a title to go to the Title Display, or use the menu at the bottom of the screen to start a new search.

Figure 3: Search Results

The Title Display (Figure 4) shows information on the selected title. Also, from Title Display the user may click on the Author or Subject to perform a search based on that term. This feature is known as a hyperlink search.

Figure 4: Title Display



Current Limitations of the OPAC

Currently users cannot limit searches by date or language, sort their search results, or combine criteria from multiple indexes. Support for boolean searching is also not available at this time.

The Transaction Log

The transaction log of this OPAC keeps a record of each time the user makes a request of the system; this roughly translates into every time the user switches screens. The fields kept include type of transaction, date, time, session number, search string, record number (NumRec), title identifier, display type, index searched, the screen requested (ToScr), the current screen (FromScr), and type of search (Key). These data elements are described in Table 1 below. As the location of the user is not passed to the system during a request, this information is not stored in the log.

Table 1: Data Elements in the Transaction Log

Data Elements	Description
Type	The type of transaction, described in Table 2 below.
Date	The date the transaction occurred
Time	The time the transaction occurred. Stored as the number of seconds past midnight.
Session	The session number. This is incremented each time the “New Search” link is followed or the time out function is activated. At CCC&TI the time out function is set to activate after XX minutes.
SearchStr	The string of text that was searched on.
NumRec	Depending on the transaction type, this field represents either the number of records retrieved by the search or the number of the record that was viewed.
Title	The database ID of the title that was viewed, if available.
Display	The type of display requested. Available display formats include Table, Long, and Short. This feature is available mostly for use with adaptive workstations and is not used at CCC&TI.
Index	The index that was searched.
ToScr	The screen that was being requested.
FromScr	The screen that the user was currently on.
Key	Whether the search was performed as a keyword search or not.

The transaction types include those listed in Table 2 below.

Table 2: Transaction Types in the OPAC Log

Transaction Type	Description
Initialize	This transaction type is registered when either the time out feature engages or the user clicks on “New Search.” This function clears all search criteria and increments the session counter.
NewSearch	Registered when a user requests a search from the search screen after an initialize has taken place.
ExecSearch	Registered whenever a search is performed. NumRec shows the number of hits retrieved.
ModifySearch	Registered when the user requests a search after having clicked on “Modify Search.” Modify search brings up the search screen with the fields filled in with the criteria of the last search performed.
Search	Registered when the user requests a search from the Title Display (Hyperlink searching) or from the Browse.
NextResults	Registered when a user requests the next page of results. NumRec shows the number of the highest record shown on the requested page.
PrevResults	Registered when a user requests the previous page of results. NumRec shows the number of the highest record shown on the requested page.
DispTitle	Registered when a user requests to view a title from the search results screen. NumRec shows the record number of the title to be displayed and Title shows the system ID of the record requested.
NextTitle	Registered when the user requests the next title to be displayed from the Title Display. NumRec shows the record number of the title to be displayed.
PrevTitle	Registered when the user requests the previous title to be displayed from the Title Display. NumRec shows the record number of the title to be displayed.
NewBrowse	Registered when the user requests a browse. NumRec shows the number of entries displayed (always 25).
NextBrowse	Registered when the user requests the next page of a browse list. NumRec shows the number of entries displayed (always 25).
PrevBrowse	Registered when the user requests the previous page of a browse list. NumRec shows the number of entries displayed (always 25).

A sample of the transaction logs is shown in Figure 5 below.

Figure 5: Sample of the Transaction Log Data

Type	Date	Time	Sessi	SearchStr	NumRec	Index	ToScr	FromScr	Key	Reason
DispTitle	2/10/99	67040	876	Self esteem in children	4	AnySubject	Title	SRES	no	
DispTitle	2/10/99	67040	875	Early childhood	1	AnySubject	Title	SRES	no	
Search	2/10/99	67061	875	Parenting		AnySubject	SRes	Title	no	
ExecSearch	2/10/99	67062	875	Parenting	122	AnySubject	SearchRes	SearchRes	no	
NewSearch	2/10/99	67073	878			AnySubject	NewSearch	SRES	no	
NewSearch	2/10/99	67075	879			AnySubject	NewSearch	Title	no	
NewSearch	2/10/99	67090	876	parenting education		AnySubject	SRes	SimpleSearch	no	
ExecSearch	2/10/99	67090	876	parenting education	0	AnySubject	SearchRes	SearchRes	no	c
NewSearch	2/10/99	67098	880			AnySubject	NewSearch	SRES	no	
DispTitle	2/10/99	67102	876	Discipline of children	15	AnySubject	Title	SRES	no	
NewSearch	2/10/99	67109	875	Attention deficit disorder		AnySubject	SRes	SimpleSearch	yes	
ExecSearch	2/10/99	67109	875	Attention deficit disorder	16	AnySubject	SearchRes	SearchRes	yes	
Search	2/10/99	67110	866	Manipulative behavior		AnySubject	SRes	Title	no	
ExecSearch	2/10/99	67110	866	Manipulative behavior	6	AnySubject	SearchRes	SearchRes	no	
NewSearch	2/10/99	67117	866	parenting education		AnySubject	SRes	SimpleSearch	no	
ExecSearch	2/10/99	67117	866	parenting education	0	AnySubject	SearchRes	SearchRes	no	c
NewSearch	2/10/99	67122	881			AnySubject	NewSearch	SRES	no	
NewSearch	2/10/99	67139	876	parenting education		AnySubject	SRes	SimpleSearch	yes	
ExecSearch	2/10/99	67139	876	parenting education	0	AnySubject	SearchRes	SearchRes	yes	c
Search	2/10/99	67143	876	Corporal punishment		AnySubject	SRes	Title	no	
Search	2/10/99	67143	876	Corporal punishment		AnySubject	SRes	Title	no	
ExecSearch	2/10/99	67143	876	Corporal punishment	1	AnySubject	SearchRes	SearchRes	no	
ExecSearch	2/10/99	67143	876	Corporal punishment	1	AnySubject	SearchRes	SearchRes	no	
DispTitle	2/10/99	67144	866	Manipulative behavior	4	AnySubject	Title	SRES	no	

Methodology

Overview

Transaction logs were collected for the period from February 8 to March 8, 1999, gathering a total of 5,479 searches. Two methods were used to analyze the logs, a series of automated analyses and a detailed manual analysis of zero-hit searches.

Data Collection

In traditional transaction log studies the number of transactions occurring in even a week would be overwhelming to review. Therefore the researchers often pared down the amount of data they obtained by collecting data from a limited time period and / or limiting the number of workstations used. For example Hunter (1991) looked at 13 hours of transactions and had over four thousand searches for the first round of evaluation. Wallace (1993) looked at 20 hours of transactions from 11 of 82 terminals and also obtained over four thousand searches to analyze. A table comparing the sample sizes from other transaction log studies can be found in the discussion section on page 43.

This study was conducted on a library that is considerably smaller than any of the published studies reviewed. As such, in order to get a sufficiently large sample a longer time period was used. For this study the transaction logs from the period February 8, 1999 to March 8, 1999 were obtained. This included 19,186 transactions, of which 5,479 were executed searches. Since this figure was comparable in size to

the samples selected for studies done at larger academic libraries (Hunter 1991, Wallace 1993, Zink 1991), all of the downloaded logs were used in the analysis. The time period encompassed a wide range of activity levels and included normal operations, weekends, and one day of low activity due to inclement weather.

Since the OPAC is almost exclusively used in the LRC itself, the searches were, for the most part, performed only during operating hours which are Monday - Thursday 8am - 9pm, Friday 8am - 4:30pm, and Saturday 8am - 2pm. There were 82 searches performed outside of operating hours; these searches are likely done by library staff or faculty members who have access via the campus network. The characteristics of these searches seemed to be similar to those of the searches as a whole. As such they were not excluded from the analysis. At this time there is no way to tell how many of the searches were done by faculty members outside of the library during operating hours.

Automated Analyses of the Logs

In KLAS the transaction logs are stored in a table in the relational database used to manage the system. For this study, several automated database queries were designed in the Progress 4GL to run against the transaction log. Progress is the relational database language that the KLAS system is written in. One query identified the individual searches.

Number of Searches Executed. This query was meant to present a summary look at the number of searches performed against the database and an idea of how many of the searches were unsuccessful. To gather the data the program looks at transactions with the type “Exec-

Search" and accumulates counts by index for total searches and searches with zero hits returned.

The next set of analyses was designed to look at usage patterns over time to determine if and when periods of high usage occurred.

Searches by Date. Transactions are sorted by date and the number of ExecSearches are counted for each date range (each day during the month of data collection). The figures are accumulated for the total number of searches and the number of zero hit searches.

Searches by Day of Week. Transactions are sorted by date and the number of ExecSearches are counted for each day of the week. The figures are accumulated for the total number of searches and the number of zero hit searches. For example, the number of searches conducted on all the Mondays during the data collection period was calculated.

Searches by Time of Day. Transactions are sorted by time and the number of ExecSearches are counted for each hour that a search was conducted. The figures are accumulated for the total number of searches and the number of zero hit searches.

This analysis was first run for the month as a whole. However, since there was so much variation in the number of searches by day of the week, this analysis was run for each individual date and figures were compiled for each day of the week using a spreadsheet. Giving a count, for example, of the number of the average number of searches conducted between 9:00am and 10:00am on Mondays during the data collection period.

The next set of queries was designed to look at how often problems occurred in the searches.

Number of Titles Retrieved per Search. In addition to zero hit searches, searches that return lots of records can be frustrating. This query was meant to get an idea of how the number of results returned were distributed. It looks at transactions with the type "ExecSearch" and accumulates counts of the number of searches retrieving sets of each possible size. The outcome of this analysis was further organized by the index used.

Searches after zero hit searches. This analysis looks at the extent to which users are able to recover from a zero hit search. It scans through transactions of type "ExecSearch" sorted by session, date, and time. When it finds a search with zero hits it looks at the next search unless it is from a new session and records the number of titles retrieved.

Retrieval rate for re-executed searches. The purpose of this query was to determine the possible effect of adding to the OPAC the capability to search all indexes simultaneously. For each transaction with the type "ExecSearch" three new searches were run using the same search string and keyword / begins setting. The searches were run against the each of the other two indices and then against a combination of the three indices.

Zero hit searches found in an all-index search. This query is meant to look more specifically at the effects of searching all indices on the zero hit subject searches. It went through all the transactions of type "ExecSearch," index "AnySubject," and numRec "0." Then the query performed each search against the Title and General indices with the same search string and keyword/begins setting. The number of searches were accumulated by the size of the set retrieved with the new searches.

The next set of queries was designed to look at how often users were taking advantage of features other than the basic search functions.

Number of Titles Viewed per Search. A measure of user interest in a search can be the number of titles in the search for which they request more information. (This measure can be problematic because the user may recognize a desired title from the search results screen and just write down the call number without viewing more information.) In order to compile this information, transactions were sorted by session, date, and time. The file was then sequentially read. When an execute search transaction was encountered a flag was set and the number of times the title display appeared was counted until there was a new search executed or the session number changed. The number of searches were cumulated by how many title displays were counted and by index.

Number of Hyperlink Searches. The next query was meant to show if and how the hyperlink search feature was being used. Hyperlink searches allow the user to perform a new search directly from the title display by clicking on the author or title. To count the usage of this feature the program looked at transactions for type "Search" or NewSearch" and FromScr of "Title." The number of hyperlink searches was accumulated by index.

Number of Browses. The final query was designed to show if an how the browse feature was being used. The browse feature allows users to scan through an alphabetical list of index terms for the chosen index. To count the usage of this feature, the program looked at transactions with the type "NewBrowse". The number of browses was accumulated by index.

Detailed Analysis of Zero-Hit Searches

The transaction logs were also imported into a Microsoft Access table for visual review and to assign probable reasons for the cause of zero hit searches. Author searches were also run against the University of North Carolina catalog to see if the search was for a valid author. Seven categories of possible errors were defined:

Misunderstood the system. These searches suggested a gross misunderstanding of how controlled vocabularies and/or KLAS works. They included search strings with natural language-like input, searches with five or more non-stop word terms, and searches following zero hit searches that added terms. Examples of these entries include: "college athletes leaving early for the pros," "right and left brain oriented math students," and "treatments for spinal cord injury" following a search for "spinal cord thearpy."

Trouble with Controlled Vocabulary. These searches seemed to be caused by a failure of the controlled vocabulary to match the words that the users were actually entering. Some examples (term used / LCSH term) of such problems include: ultrasound / Ultrasonography, insanity plea / Insanity - Jurisprudence, and personal computers / Microcomputers.

Spelling and Typographical Errors. This category included searches with one or more of the terms misspelled, typographical errors, and junk entries. Examples of these searches include "Emily Dickinson," "medievil," "comotology," and "asdfjlk' asdf jkl;' asdf jl;' asdf kl;."

Searched the Wrong Index. This category was used for searches where clearly the wrong index was selected. Examples of this included "Green Eggs and Ham" as a subject search, "asthma" as an author search, and "Ayer, Eleanor" as a title search.

Keyword v. Begins Problem. These searches seemed to indicate either that the user used "begins" when a keyword search would have been more appropriate, or tried to search on truncated words with a keyword search.

Collection Failure. There was no apparent reason for these searches to fail, suggesting that the item was not in the collection. This category included authors that were found in the UNC catalog but not in the Caldwell database, specific subject searches for which the titles available were more general, and title searches that seemed otherwise correct.

Unknown Author. For these searches, it was not possible to find an author entry in the UNC catalog with the search string given. Examples of these searches include "twaynes," "Leeds, Mary," and "cleage."

To further understand the dynamics of author searching against this database, all of the author searches were given an additional classification. The searches were classified according to the main publishing area for which the author is known. For example, John Grisham and Sharyn McCrumb were classified as fiction, Bethany Campbell a juvenile literature, Bob Dylan as popular culture, Emily Dickinson and Edgar Allen Poe as literature/poetry, Thomas Jefferson as historical, David Abrahamsen as social science, and Edward Teller as science. Authors who were not findable in the UNC catalog and search strings that were not author like were classified as unknown/error, for example, "asthma" as an author search was an error.

Results Analysis

Usage Patterns

The total number of searches for the 28 day period was 5,479 searches, averaging just under 230 searches per day the library was open. The vast majority of searches were on the subject index at over 83% of the total searches performed. As a whole the searches returned zero hits about 35% of the time. The figures shown represent both keyword and begins searches. The results are detailed in Table 3 below.

Table 3: Searches Executed

Index	Searches	% of Total	Zero Hit Searches	% Zero Hit
Author	311	5.7%	95	30.5%
Subject	4,564	83.3%	1,655	36.3%
Title	604	11.0%	196	32.5%
Totals	5479	100.0%	1946	35.5%

Next, the activity was analyzed to determine any usage peaks or patterns over the time period. The usage of the library showed a fair amount of variability by date. The second week appeared to have a particularly heavy load, and February 24 was particularly slow due to inclement weather. These figures are shown in Figure 6 and Table 4 below.

Figure 6: Graph of Usage by Date

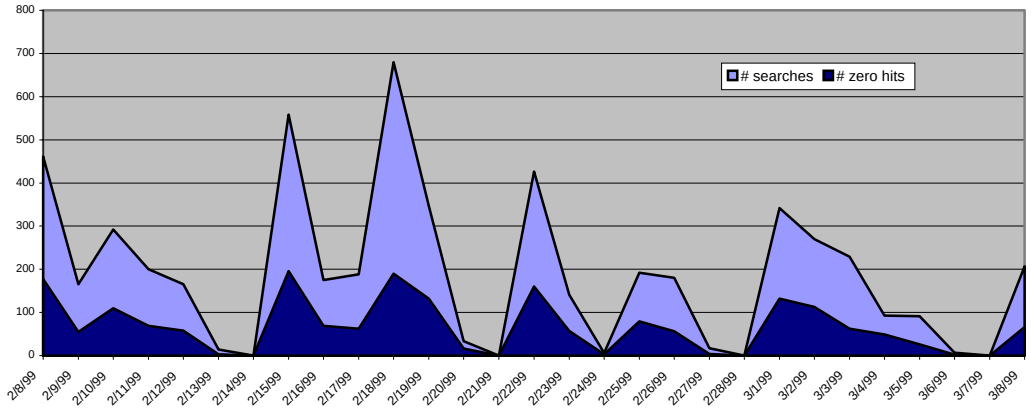


Table 4: Usage by Date

Date	# searches	% of Total	# zero hits	% zero hit	Date	# searches	% of Total	# zero hits	% zero hit
2/8/99	461	8.41	177	38.39	2/22/99	426	7.78	160	37.56
2/9/99	165	3.01	55	33.33	2/23/99	141	2.57	57	40.43
2/10/99	292	5.33	110	37.67	2/24/99	6	0.11	3	50.00
2/11/99	200	3.65	69	34.50	2/25/99	192	3.5	79	41.15
2/12/99	165	3.01	58	35.15	2/26/99	180	3.29	56	31.11
2/13/99	14	0.26	3	21.43	2/27/99	17	0.31	4	23.53
2/14/99	0		0		2/28/99	0		0	
2/15/99	558	10.18	196	35.13	3/1/99	342	6.24	132	38.60
2/16/99	175	3.19	69	39.43	3/2/99	270	4.93	113	41.85
2/17/99	188	3.43	62	32.98	3/3/99	229	4.18	62	27.07
2/18/99	680	12.41	190	27.94	3/4/99	93	1.7	49	52.69
2/19/99	347	6.33	132	38.04	3/5/99	91	1.66	26	28.57
2/20/99	33	0.6	16	48.48	3/6/99	7	0.13	2	28.57
2/21/99	0		0		3/7/99	0		0	
					3/8/99	207	3.78	66	31.88

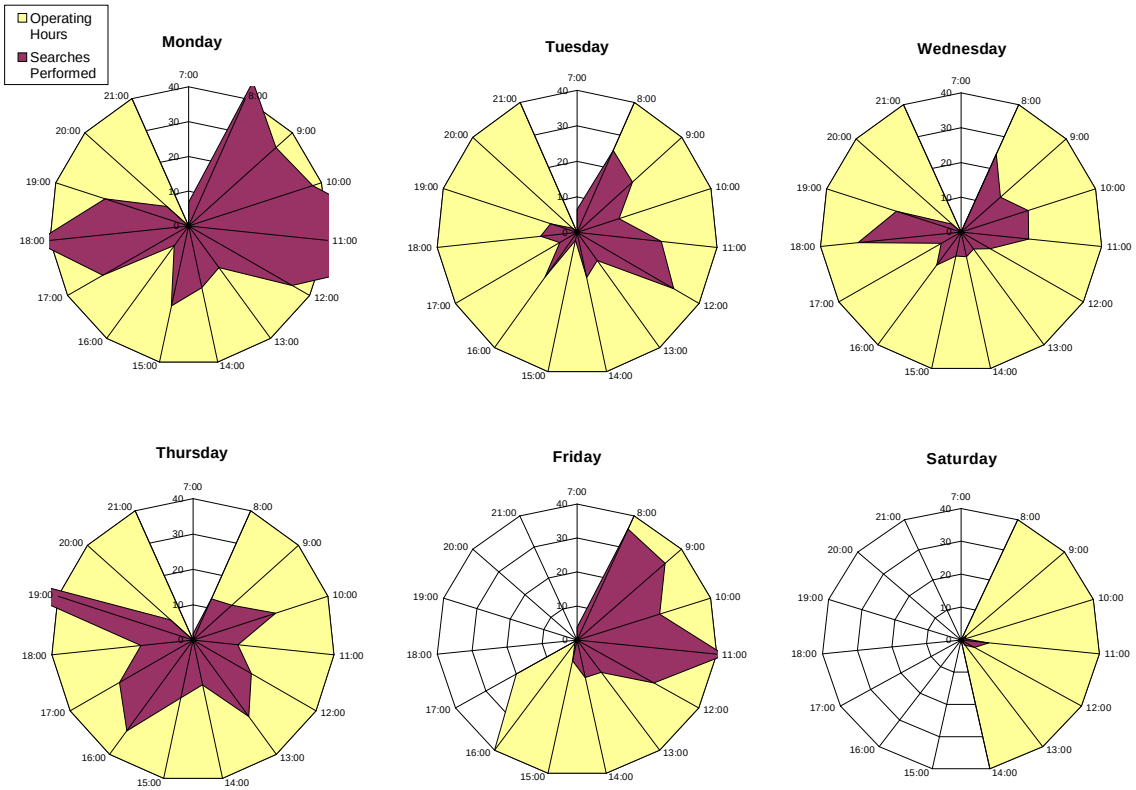
In searches arranged by day of the week Monday had the heaviest usage by far, with an average of almost 400 searches per day. Thursday also had high activity with almost 300 searches. The library catalog was rarely used on Saturdays, even though the library was open for six hours. The figures for this analysis are shown in Table 5 below.

Table 5: Usage by Day of the week

Day of week	Total # searches	Avg # Searches	% of Total	# zero hits	% zero hit
Monday	1,994	398.8	36.39	731	36.66
Tuesday	751	187.8	13.71	294	39.15
Wednesday	715	178.8	13.05	237	33.15
Thursday	1,165	291.3	21.26	387	33.22
Friday	783	195.8	14.29	272	34.74
Saturday	71	17.8	1.3	25	35.21

In order to further see the trends in usage the number of searches were counted for each hour of each day. These figures were combined by day of the week and tabulated in a series of graphs shown in Figure 7 below. For each day the graph shows the average number of searches in a given hour over the month of data collection. Notable spikes occur on Monday mornings and evenings and Thursday evenings.

Figure 7: Usage graphed by time of day for each day of the week



Search Failure Analysis

The number of results retrieved by each search were tabulated to see how many searches were retrieving zero hits or unrealistically large numbers of hits (Table 6). Zero hit searches represented 30 - 35% of executed searches. Searches retrieving more than 100 hits accounted for about 15% of the searches ranging from under 1% in the author index to 32% in the title index.

Table 6: Number of Titles Retrieved

Index	0	1 - 5	6 - 10	11 - 25	26 - 50	51 - 100	100 +
Author	95	82	17	68	42	6	1
Subject	1655	715	289	371	394	373	760
Title	196	219	30	65	19	22	53
Totals	1946	1016	336	504	455	401	814

Once a user retrieves a zero hit search, their ability to identify the problem and enter a new search that will retrieve hits is of great importance. In this study users were able to retrieve at least one hit on over 70% of the subsequent searches. Those searches averaged about one page worth of results. See Table 7 for details.

Table 7: Searches performed after zero hit searches

Index	# zero hit Searches	# zero hit on Subsequent	% Corrected	Avg Hits on Subsequent
Author	95	31	67.4%	23.45
Subject	1,655	431	74.0%	25.57
Title	196	50	74.5%	9.91
Totals	1946	512	73.7%	24.05

In classifying the reasons why users got zero hits, the most prevalent problem was a lack of understanding about how controlled vocabularies work, comprising just over 60% of the zero-hit searches. Typographical errors and spelling mistakes accounted for 15% of the errors, many of these errors appeared to have been corrected in subsequent searches. Another prevalent user error was misunderstanding the structure or scope of the system such by adding terms to the search string after

a zero hit search or entering searches that would have been more appropriate to a magazine index. The final major issue were search terms that may have been found in other library collections but were not in the CCC&TI collection.

Table 8: Reasons for Zero-hit Results

Reason	# Searches	% of Total
Misunderstood the system	163	8.4%
Spelling / Typographical Error	292	15.0%
Trouble with controlled vocabulary	1171	60.2%
Searched the wrong index	68	3.5%
Keyword v. Begins Problem	38	2.0%
Not in the system	203	10.4%
Unknown Author	11	0.6%
Total	1946	

Knowing that understanding the controlled vocabulary was a likely problem, it became important to see what kinds of things could be done to fix the problem. One solution that could be implemented relatively quickly is to add "all indexes" as a search option, which would allow a keyword search of uncontrolled vocabulary terms such as those in the title. The concern with this approach is that it will greatly increase the number of searches returning too many hits. In order to determine the impact of such a step the searches were run against the database again as an all indexes search. The results, as described below, were generally positive.

For author searches, the zero hit rate dropped from 30% to 27%, and the percentage of searches with more than 50 hits rose from 2% to 10%. For title searches, there was almost no change in the zero hit rate or the percentage of searches with more than 50 hits. For subject searches, the zero hit rate dropped from 36% to 26%, and the percentage of searches with more than 50 hits rose from 25% to 29%. Subject searches were also compared against the effectiveness of using the title index

in lieu of the controlled vocabulary. This strategy appeared to cancel out successes and failures (Figure 9). The overall effect of performing all of the searches as all-index searches decreased the zero hit rate from 36% to 27% and increased the percentage of searches with more than 50 hits from 22% to 26%. The results can bee seen graphically in Figures 8, 9, 10, and 11.

Figure 8: Change in titles retrieved for searches that were originally conducted in the author index

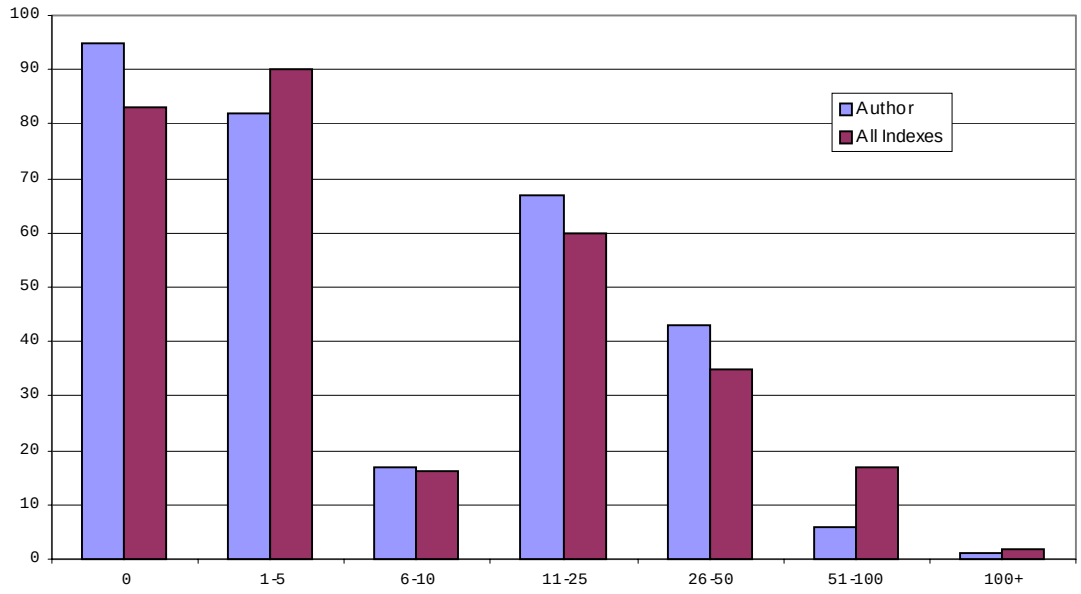


Figure 9: Change in titles retrieved for searches that were originally conducted in the title index

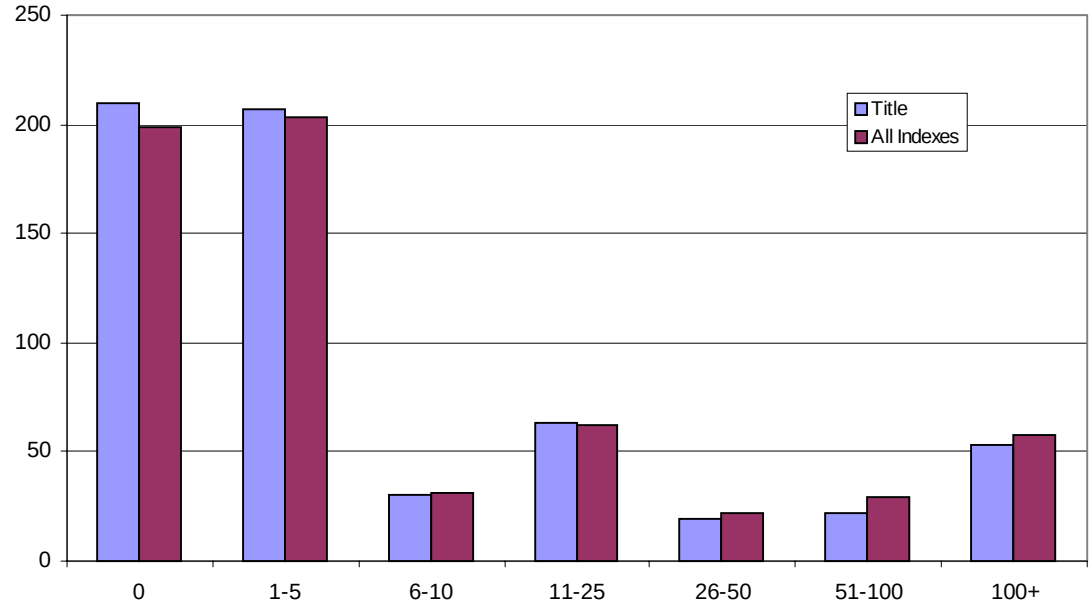


Figure 10: Change in titles retrieved for searches that were originally conducted in the subject index

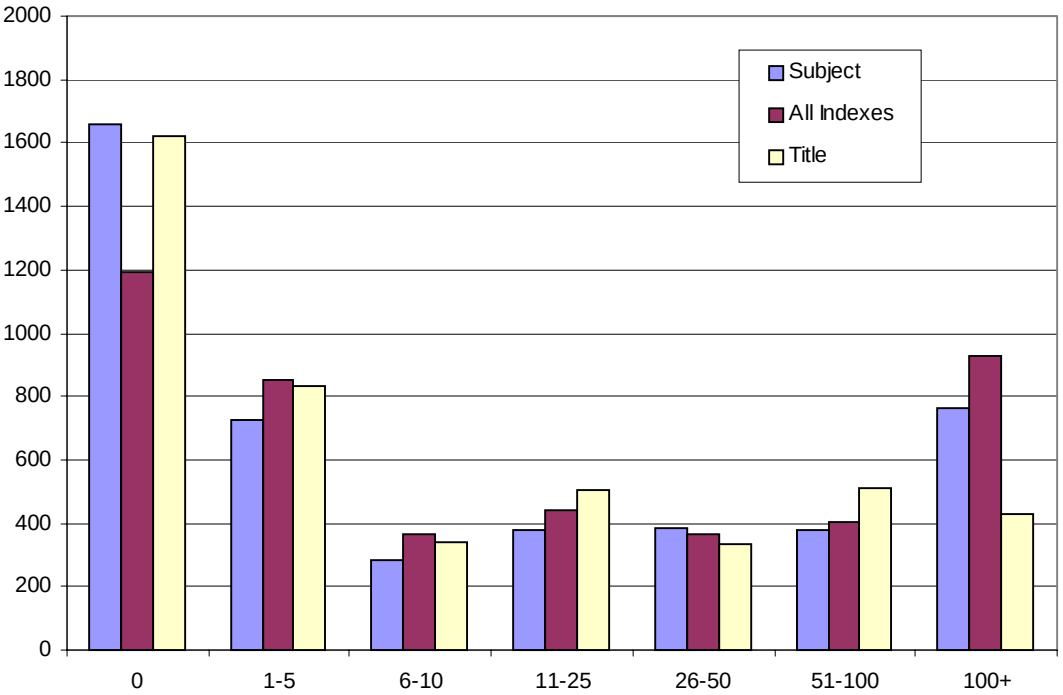
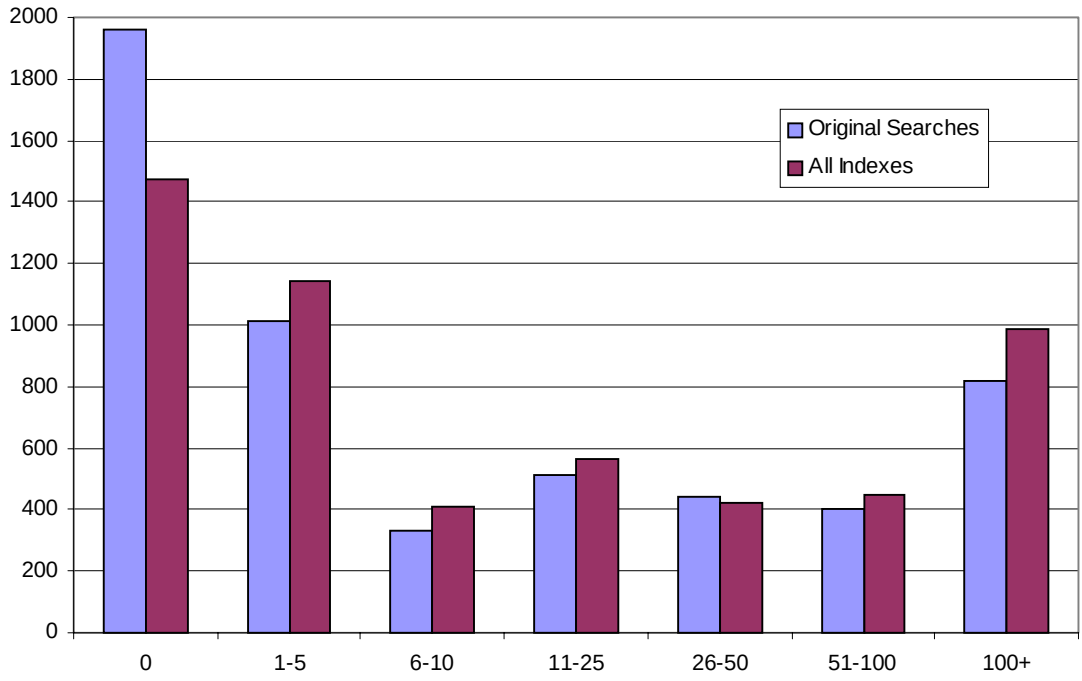


Figure 11: Change in titles retrieved for all original searches regardless of index.



When only the zero-hit searches were re-run the impact could more clearly be seen. Running the zero-hit searches against the same index shows that about 1% of the searches would be affected just by the time difference in when the search was run. However, over 25% of the zero hit searches would have returned some thing if they had been run as an all-index search. The majority of improvement in the subject index came from re-run searches in the title index with 466 searches returning non-zero results. See Table 9 for details.

Table 9: Improvements in Zero-Hit searches

Index	# zero hit searches	# searches aided by time	% aided by time	# searches aided by all-index	% aided by all-index
Author	95	1	1.1%	13	13.7%
Title	196	4	2.0%	15	7.7%
Subject	1655	13	0.8%	471	28.5%
Totals	1946	18	0.9%	499	25.6%

Usage of Other Features

The number of titles viewed for a given search were tabulated to see how interested the users were in viewing the results (Table 10). Title display were viewed for approximately 20% of the searches. The numbers from this table could be taken as an indication of how many searches users thought useful enough to examine further. This measure is problematic, however, due to the fact that the call number for a given record appears on the search results screen and therefore a search that retrieved a useful item could be missed if all the user did was to write down the desired call number and move on.

Table 10: Titles Viewed

Index	% at least 1	1	2	3-5	6-10	11-15	16-30	31-50	51+
Author	22.2%	48	9	9	3	0	0	0	0
Subject	20.6%	651	104	117	44	15	8	2	0
Title	18.9%	65	16	21	8	1	2	1	0
Totals	20.5%	764	129	147	55	16	10	3	0

The usage of additional search features was also analyzed. From the title display users can click on a hyperlink to perform a new author or subject search. The analysis (see Table 11) indicated that usage of this feature comprised about 6.5% of author and subject searches, and occurred over 12% of the time when a title from a subject search was displayed.

Table 11: Usage of Hyperlink Searches

Index	# HyperLink Searches	% of total searches	% of titles viewed
Author	21	6.8%	30.4%
Subject	285	6.2%	30.3%
Title	1	0.2%	0.9%
Totals	307	5.6%	27.3%

Browses of the indices were used much less frequently, with subject browses being the most common at 2.6% of the total number of subject searches performed. Table 12 shows the number of times the browse feature was used.

Table 12: Usage of Browse

Index	# Times Browsed	% of total searches
Author	2	0.6%
Subject	118	2.6%
Title	9	1.5%
Totals	129	2.4%

Discussion

The focus of this study was to look at patterns of searching and search failure in a small library so that areas for improvement and tools needed for further monitoring could be identified. It is somewhat difficult to generalize the results of this study due to the system specific nature of computer-human interactions and the dearth of information about searching at libraries serving similar populations. However, hopefully this study will serve as a starting point for comparing future improvements to the system and other similarly sized libraries.

System Usage Patterns

The evening spikes on Mondays and Thursdays likely represent searching done by evening classes that come to the library to develop their searching and research skills (Evert, personal communication, May 1, 1999). Currently, issues such as library hours and weather greatly affect the catalog usage (see Figure 6 and Table 4). It will be interesting to see if and how the impact of these issues changes as the catalog is made available from outside campus and more people begin to use it at off-peak times. A more personal knowledge of activities in the library and on campus would likely aid in the understanding of peaks in usage. These reports are probably most useful to the library manager who wants them as quantitative information on an aspect of library usage.

Search Characteristics

Of the 5479 searches studied the vast majority of the searches were executed as subject searches (4,564 or 83.3%). This is considerably higher than any of the other studies (see Table 13). The figures for author and title were correspondingly lower than other studies at 5.7% and 11.0% of the searches respectively.

Table 13: Sample Data from other Transaction Log Studies

	Peters, 1989	**Hunter, 1991**	**Zink, 1991**	**Wallace, 1993**	**Wyly, 1996**	**Atlas, 1997**	**Holloman, 1999**
# Searches Analyzed	9,565	3,707	6,118	4,134	795,810	–	5,479
Author / Name	22.3% (30.2%)	21.4% (42.3%)	13.4% (30.6%)	21.7%	13% (22%)	18%/14%	5.7% (30.5%)
Title	34.3% (43.8%)	25.5% (47.9%)	19.3% (43.0%)	24.2%	25% (32%)	17%/24%	11.0% (32.5%)
Subject	31.9% (0%-browse)	51.8% (62.0%)	49.3% (22.1%)	–	26% (70%)	43%/43%	83.3% (36.3%)
Keyword	–	–	–	53.1%	–	22%/19%	–
Average zero hit rate	27.9%	54.2%	27.8%	10.4%	44%	–	35.5%

Note: Numbers in parentheses indicate the percentage of searches in that index that returned zero hits.

The high percentage of subject searches seems to go against Larson's (1991) observations of a decline in subject searching. Potential reasons for the predominance of subject searches in this study reflect several factors. First, the default setting for searches is a subject search; users must actively choose to use an index other than subject. Second, keyword searches -- especially on the title field -- have been gaining popularity in the larger systems because they allow topical searches without having to depend on the controlled vocabulary. While the KLAS OPAC allows for keyword searching from any index, it does not currently allow for the

combination of indices which cover both titles and subjects. A third factor might be the educational level of the population using the catalog. As Ferl (1996) noted, users such as graduate students and university faculty members with a high familiarity in a specialized field will be much more likely to search from a list of known authors or titles. However, in more generalized areas of study, a subject is the "known" facet of the search. Finally, the problem of information overload that Larson (1991) suggested might drive users away from subject searching is not highly present in this catalog. Few searches will return "too many" hits in a database of 48,500 titles.

It should also be noted that the author searches were just half of the number of title searches, which is substantially lower than the figures for other studies. Zink reports 13.35% author and 19.32% title; Hunter reports 25.5% author and 21.4% title; and Atlas reports 18% author and 17% title in the first phase of his study with 14% author and 24% title in the second phase of his study.

Also, based on the visual analysis of the logs, certain authors appeared frequently and the majority of those were fiction and literary authors such as Grisham, Dickinson, and Poe. Table 14, shown below, indicates the distribution of author searches based on the field in which the author generally publishes.

Table 14: Author Searches organized by publishing field of the author

Publishing Field	Number of Searches	% of Total Author Searches
Fiction	121	38.9%
Unknown / Error	62	19.9%
Literature / Poetry	38	12.2%
Social Science	35	11.2%
Popular Culture	13	4.2%
Science	10	3.2%
Humanities	10	3.2%
Historical Figure	7	2.2%
Autobiography	6	1.9%
Art / Photography	5	1.6%
Juvenile Literature	4	1.2%

It also appears that exposure to the internet has had some effect on users' searching strategies. Although using the OPAC is closely related to the traditional library setting, an increasing number of the other resources used in the library, such as NCLive, are also accessed through the web browser. The effect of internet search engines can mostly be seen in the users' syntax. Fifty-four searches, representing 31 unique search strings, used "+" in the search string. Long search strings were another common feature; 233 subject searches had more than 25 characters in the search string, 14 had more than 50 characters. Further some users would add more terms to their search strings after receiving zero-hit results in an attempt to expand their search.

There were several search strings that appeared multiple times. Overall there were 70 search strings that were entered 10 or more times. “Opposing viewpoints” topped the list with 52 searches, and “computers” came in second with 50 searches. The Opposing Viewpoints Series is put out by Greenhaven Press and covers a wide range of ethical, political, and other controversial issues. At the time of this study the library had 304 titles from this series. One of the problems in finding this series is that the user needs to look in the title index instead of the subject index. The most common zero-hit searches were “air bags” with 10 entries and “history of math” with 8 entries.

In KLAS if a search retrieves 200 hits the user is asked if they would like to continue with the search, view the first 200, or cancel the search. In total 434 searches reached this number, serving around 50 unique search strings. Of those searches 288 stopped at 200 records. The largest search completed was “history” with 6070 records retrieved; the search was requested 3 times. Other common large searches included “music” with 639 hits, requested 7 times; “nursing” with 496 hit, requested 17 times; and “sex” with 282 hits, requested 20 times.

Search Failures

The percentage of zero hit searches was between 30% and 36% which is both better and worse than other studies. Hunter reports a low of 42.3% for author and a high of 62.0% for subject with an average of 54.2%. Zink reports a low of 22.1% for subjects and a high of 41.03% for titles with an average of 22.82%. Wallace reports an average of 10.4% zero hit searches.

The greatest concern about the zero hit searches is that over 60% of them are coming from problems with the controlled vocabulary. However, the potential for improvement seems high. Although topic searching by title is by no means a silver bullet, the results of this study (Table 9 and Figure 10) seem to support Gerhan's (1989) observation that for the number of titles that are topically accessible by their titles, there are also a substantial number that require subject enhancement; as many as 25% of the zero-hit subject searches should find at least one record if changed to an all indexes search. Further, studies such as Wilkes (1995) suggest that over 70% of zero-hit searches can find at least one hit if links are provided for "see" references.

Feature Usage

In her study of the University of Colorado system Wallace (1993) evaluated the use of system supplied search aids. Of the 11 search aids evaluated only two were used in more than 2% of the searches. They were Quick Search (8.7%), which allows the user to shorten the number of keystrokes for a search to three plus the search string, and Search History (3.2%), which allows the user to review the most recent searches. Another search aid that was available on the CU system was Express Search, which as described is similar to the hyperlink searching available in the KLAS OPAC. This feature, however, was only used 14 times out of 4,134 searches (0.3%) during Wallace's study. Wallace gives a couple of suggestions for possible reasons that the usage of these features was so low. The first is that there was not much indication on the screen itself that a particular feature was available. Second was that the description of these features was relatively buried in the help

system and as such it was quite unlikely that the users were aware of their existence. Further in the case of the express search, she felt that the situations in which the feature would be desirable were limited.

In the current, the results indicate that the affordances for hyperlink searches seem to be adequate as about 30% of the users going to the title display for author and subject searches chose to follow the links. It is probable that exposure to the internet has increased the visibility and understanding of this feature.

Browses, on the other hand, need considerably more visibility. The first area that needs to be addressed is the placement of the browse buttons on the search screen. Due to the settings that the library uses for NCLive, the submit and browse buttons are scrolled off the first screen and are not visible to the user. Further it appears to be unclear to some users how the browse should be used, they enter nothing in the search box and begin scrolling through the subjects starting with "35mm cameras" to "Abdomen Ultrasonic Imaging." After about 15 screens or so they get fed up and start over. An idea is to provide a link into the approximate alphabetical section, as a suggestion for users having difficulty finding matches more directly.

Suggestions for Improvement

The first two suggestions for improvements have to do with the indexing of the catalog. First, an all-indexes option needs to be provided. The addition of such a feature may lower the number of zero-hit searches for the predominant type of search by as much as 25%. Second is to implement authority control. A data entry and searching mechanism needs to be devloped by the system, and the library

needs to develop the appropriate see and see also references. This improvement requires considerably more work on the part of both the library and the system designers but is the best long term solution for dealing with the controlled vocabulary issues.

Feedback needs to be provided to the user upon retrieving a zero hit search. A table showing the number of hits each term in the query gets by itself would serve to show the user that the system is looking for all of the search terms and what their most restrictive terms are. Such a display may also help to point out spelling or typographical mistakes. Another useful piece of assistance on the feedback screen may be a link to the alphabetical spot in the browse list for each of the terms entered. This would let the user go to approximations of a word if they are unsure of the spelling.

The final set of suggestions deals with processing the entered search string to achieve more successful matching with the database records. First, stop words such as “and,” “or,” “of,” “the,” “in,” “as,” “a,” and “an,” need to be stripped from the search strings in keyword searches. Next a strategy for dealing with apostrophe’s, needs to be developed. Finally a stemming and/or “spell check” algorithm should offer other possible words that appear in the database. The options should appear on a feedback page. Each of these suggestions should lower the number of zero-hit searches.

Tools needed for Management Reports

The kinds of information that need to be gathered and organized for easy evaluation on a regular basis include: system usage levels, search failure rates, feature usage levels, and sample searches. Reports on system usage levels, over all and by time period, should be able to help a manager with justification of the system and any improvements that are made. Search failure rates provide a measure of success. Consistent reporting of these will allow the manager and designer to know if improvements are having the desired effect or if something went wrong. Feature usage levels help to identify aspects of the system that are being used and give another view of how users are taking advantage of the system. Sample searches such as lists of common zero-hit searches, common successful searches, and searches retrieving too many hits, give further suggestions on what terms users are actually using to search, and may be used for suggestions on additional entries in the catalog or additional items in the collection.

The addition of location information in the transaction logs would allow each of the types of reports described above to be run for individual user groups (users at the main campus, users at the Boone campus, users from off campus). This information would be useful in determining the extent to which a given group is using the system, and if there are any particular concerns need to be addressed with a particular group.

Conclusions

More study is needed to determine the searching behaviors of patrons in settings other than large university libraries. It seems reasonable to expect differences in experience with computers and knowledge of searching techniques from patrons in smaller and more rural libraries. As we grow more and more connected, the pool of users grows further and further away from just the academic circles surrounding large universities. We need to understand the behaviors of these outlying communities so that we can better serve them.

There are always improvements to be made in any system, and this one is no exception. Identifying problem areas and setting up monitoring tools are important steps to improving the capabilities of the system. Transaction log analysis is a tool that should be able to help this and other libraries using the system evaluate their collections, cataloging, and search capabilities.

Bibliography

Atlas, M. C.; Little, K. R.; Purcell, M. O. (1997). Flip charts at the OPAC: using transaction log analysis to judge their effectiveness. *Reference & User Services Quarterly,* 37(1) 63-9.

Baker, R. K. (1997). Faculty perceptions towards student library use in a large urban community college. *The Journal of Academic Librarianship*, 23(5) 177-182.

Blecic, D. D.; Bangalore, N. S.; Dorsch, J. L. (1998). Using transaction log analysis to improve OPAC retrieval results. *College & Research Libraries,* 59, 39-50.

Borgman, C. L. (1989). All users of information-retrieval systems are not created equal - An exploration into individual-differences. *Information Processing & Management*, 25(3), 237-251.

Borgman, C.L. (1996). Why are online catalogs still hard to use? *Journal of the Ameican Society of Information Science*, 47(7), 493-503.

Caldwell Community College and Technical Institute (No Date). *Caldwell Community College and Technical Institute*. [On-line] Available: http:// www.caldwell.cc.nc.us/ [1999, April 3]

Drabenstott, K. M; Weller, M. S. (1996). Handling spelling errors in online catalog searches. *Library Resources & Technical Services*, 40(2), 113-132.

Ferl, T. E.; Millsap, L. (1996). The knuckle-cracker's dilemma: a transaction log study of OPAC subject searching. *Information Technology and Libraries,* 15(2), 81-98.

Gerhan, D. R. (1989). LCSH *in vivo*: subject searching performance and strategy in the OPAC era. *The Journal of Academic Librarianship*, 15(2), 83-89.

Hunter, R. N. (1991). Successes and failures of patrons searching the on-line catalog at a large academic library: a transaction log analysis. *RQ,* 30, 395-402.

Larson, R. R. (1991). The decline of subject searching: long term trends and patterns of index use in an online catalog. *Journal of the Ameican Society of Information Science*, 42(3), 197-215.

Matthews, J. R. (1982). *A Study of Six Online Public Access Catalogs: A Review of Findings.* Washington, DC: Council on Library Resources. (ERIC Document Reproduction Service No. 231 389).

Matthews, J. R; Lawrence, G. S. (1984). Further analysis of the CLR online catalog Project. *Information Technology and Libraries,* 3(4), 354-376.

Ortiz-Repiso, V.; Moscoso, P. (1999). Web-based OPACs: between tradition and innovation. *Information Technology and Libraries,* 18(2), 68-77.

Peters, T. A. (1989). When smart people fail: an analysis of the transaction log of an online public access catalog. *The Journal of Academic Librarianship*, 15, 267-73.

Peters, T. A. (1993). The History and Development of Transaction Log Analysis. *Library Hi Tech,* 11(2), 41-66.

Peters, T. A. (1996). Using transaction log analysis for library management information. (presented at the 1995 ALA Conference) *Library Administration & Management,* 10, 20-25.

Sandore, B. (1993). Applying the Results of Transaction Log Analysis. *Library Hi Tech,* 11(2), 87-97.

Tenopir, C. (1997, May 1). Common End User Errors. *Library Journal*, 122, 31-32.

Wallace, P. M. (1993). How do patrons search the on-line catalog when no one's looking? Transaction log analysis and implications for bibliographic instruction and system design. *RQ,* 33, 239-252.

Wilkes, A.; Nelson, A. (1995). Subject Searching in Two Online Catalogs: Authority Control vs. Non-Authority Control. *Cataloging & Classification Quarterly*, 20(4), 57-79.

Wyly, B. J. (1996). From access points to materials: A transaction log analysis of access point value for online catalog users. *Library Resources & Technical Services,* 40(3) 211-236.

Zink, S. D. (1991). Monitoring user search success through transaction log analysis: the WolfPAC example. *Reference Services Review*, 19(1), 49-56.